

Topic (iii) - Administration and policy of statistical data confidentiality

**INFORMED CONSENT: BUZZWORD OR PANACEA**

Submitted by Statistics Netherlands<sup>1</sup>

**Invited paper**

**I. INTRODUCTION**

1. The production of official statistics has become considerably more complicated over the last decades. Many of these processes in the past started with a questionnaire and ended with a publication. There were many of such stovepipes, parallel and simultaneous. If there was a need for new official statistics to be produced, the usual reaction was to design a new questionnaire, to collect new data, and to release a new publication.

2. Nowadays, a double transformation of these processes is taking place within Statistics Netherlands (Van Bochove 1996). Apart from and in addition to its (primary) data collection efforts, Statistics Netherlands makes increasing (secondary) use of whatever administrative data is already available. All data pertaining to one type of statistical unit (such as persons or companies) are loaded into one input database.<sup>2</sup> At present, there are two of such databases under construction, one for persons (SSB or Social Statistical File) and one for companies (Microlab).<sup>3</sup> For each member of the population in this database a record consists of variables from a variety of sources. These records in turn serve as the basis for further statistical processing and publication.

3. On the output side, Statistics Netherlands is in the process of creating one output database (called StatLine, accessible on the Internet at <http://www.cbs.nl>) containing in principle all publishable aggregates. Paper publications do not contain the traditional tabulations anymore. They focus on specific themes and topics from a variety of angles, thus accentuating the interrelations between economic and social dimensions in a concrete setting. In short, there is not a direct link anymore from one statistical questionnaire to one particular tabular publication.

---

<sup>1</sup> Prepared by Joris Nobel. The views in this paper are those of the author and do not necessarily reflect the policies of Statistics Netherlands.

<sup>2</sup> Matching of these data is done on either a unique identification number, or on a combination of detailed background variables such as postal code, housenumber, date of birth, and sex in the case of social data.

<sup>3</sup> Of course, these databases have a hierarchical structure (persons, households, families) and are related to each other (people work within companies). Other, similar databases for other types of statistical units (locations) will be created in due time.

4. Apart from the StatLine database and the thematic publications, the release of microdata for research purposes has developed into a powerful statistical outlet during the last decade. A microdata file is one of several products along the production line of collecting, processing and releasing data. Microdata files are released for the major social surveys at Statistics Netherlands. As such they are still very much input oriented conceived, along the traditional production lines. This new deviates from a firm tradition of releasing aggregates only. It is both useful and legal, but nevertheless its legitimacy in the eyes of the respondents may not be self-evident.

5. The playing field for Statistics Netherlands has changed hand in hand with the methodological approach:

- a) Units (**data subjects**) contribute to statistical data collection in two roles: as respondents but also (and increasingly) as the objects of administrations. The next section will give some examples of the combination of these two roles.
- b) Since administrative data are primarily collected by **administrative agencies**, bilateral arrangements have to be worked out.
- c) **Researchers** and their organisations put pressure on government, parliament, and public opinion to make access to microdata for research as easy as possible.
- d) Legislation is available and further developed to protect the privacy of personal data. Apart from the **legislator**, the national **privacy authority** (*Registratiekamer*) needs to be mentioned here.

6. In the remainder of this paper, I analyse the consequences of (1) the matching of survey and administrative data and (2) the release of microdata for research purposes for our relation with data subjects in the playing field mentioned above.<sup>4</sup> In particular, I concentrate on the applicability and utility of the concept of “informed consent”. In the second section I give a variety of recent examples of the statistical use of the matching of administrative and survey data at Statistics Netherlands. I draw them from the area of social statistics.<sup>5</sup> It is the very variety of these practices that calls for a more formal policy. I then describe, in the third section, frames of reference from (1) professional codes and ethical guidelines of professional associations, (2) international legislation, and (3) national legislation. From a strictly legal perspective, one could argue that there are no legal objections to either the matching of survey and administrative data, or the release of microdata for research purposes. One formal impediment turns out to be the conditions under which some administrative data is released by its authorities to Statistics Netherlands. But from a broader ethical perspective, the legitimacy of record matching and of the release of microdata is far from self-evident. Do respondents really understand what statisticians will do with their data and do they agree with such use? Some Dutch experiences are summarised in the fourth section. The fifth and final section summarises my findings as of this moment.

7. It may be useful to observe here that survey response percentages in the Netherlands are well below averages for comparable surveys in other countries. The 1971 Population Census was the last survey to entail a response obligation for social statistics. In fact, one of the reasons why the Population Census Law was revoked since then, was popular resistance to the “Big Brother is watching you” perception. On the other hand, the Netherlands as a country is very much dependent on social research, modelling, and consultancy. The relative number of social scientists is high. And the importance of objective facts and research within the pillarised Dutch society is well documented. The Netherlands is probably unique in having three so-called planning agencies which assist, advise, and evaluate government policies in the economic, social-cultural and environmental areas (CPB, SCP, and RIVM). The population administration is very well developed. Nevertheless it has two particular characteristics: it is decentralised towards the six

<sup>4</sup> Jabine (1986) deals with the very same questions of planned linkages and release of microdata, within the USA context.

<sup>5</sup> Privacy legislation refers to personal data. In statistics there often is separate legislation for economic microdata. But it is obvious that in many cases personal and company data coincide, as for companies without employees.

hundred odd municipalities in the Netherlands, and the personal identification number used in it is not the same as the one used by the fiscal and the social security authorities.

## II. MATCHING DATA

8. The advantages of matching survey and administrative data are manifold (Kooiman and Nobel, 1997). Survey data can be ideally targeted to the research question at hand. But they are expensive in terms of the researchers' budgets and processing time, and respondents' time. Normally they pertain only to a small fraction of the population. They may be hampered by non-response and by inexact answers due to the constraints of, *e.g.*, human memory and social desirability. Administrative data, on the other hand, is readily available for the whole population, in principle at marginal cost. The quality of these administrations seems to vary, however, depending on the interests of those, whose data are registered, and those who use them within the administration. The registrations of the unemployment offices have a reputation of administrative pollution, whereas the Netherlands' population administration is excellent according to international standards.

9. With the combination of administrative and survey data one could in principle have the best of both worlds. Registrations may serve the role of sampling frames, which makes statistical inference, including corrections for non-response, considerably easier. From rich registrations one could differentiate and target fieldwork strategies and operations in order to optimise response. Furthermore, data may be easily added to a sample survey from administrations because respondents do not have them readily and exactly at hand.

10. If on the other hand one wants to use surveys in addition to statistical research based primarily on administrative data, three roles can be distinguished: quality control; translation of administrative measurements into research concepts; and addition of missing data, be they records or variables. The precise effect of these roles is to a large extent dependent on the correlation of the administrative data with the target variables of the survey. If this correlation is almost zero, the administrative data are of no relevance for the survey. On the other hand, if this correlation is almost unit, one can rely completely on the administrative data and does not need a separate sample survey.

11. Now let us give some examples of matching of administrative data by Statistics Netherlands.

*A sample of applications will be presented at the meeting,*

*Joris Nobel.*

It is easy and seductive to think of further applications. The use of fiscal data for routine matching of income variables with survey data is one of them. At least three practical and methodological reasons are quite compelling: survey questions on income take a lot of interviewing time; moreover they induce high item non-response; and finally there is a consistent downward bias in results.

## III. FRAMES OF REFERENCE

12. Guidance may be derived from three kinds of sources that I will deal with successively.<sup>6</sup> For statements about professional ethics we can turn to, *e.g.*, the International Statistical Institute (1985), the American Statistical Association (1989), and the American Sociological Association (1997). Official statements of international relevance have been issued by the United Nations Economic Commission for Europe (1992), the Council of Europe (1997), and the European Union (1990, 1995, and 1997). I will also briefly describe Dutch legislation (1988, 1996).

<sup>6</sup> I have no claim to completeness. Also my selection of texts may be somewhat arbitrary. Some Web addresses are given in the bibliography.

13. One should be aware of several qualifications. Codes and laws may have different levels of ambition. Jowell, the chair of the committee that drew up the ISI 1985 Declaration, made a distinction between an *aspirational*, a *regulatory*, and an *educational* code (Jowell 1981). He rejected the expression of unattainable ideas, and rules to govern behaviour based on the premise that there are universal (and enforceable) models of appropriate practice. Instead he preferred the latter kind of code that “would seek to describe and explain professional norms, expose inherent conflicts and give guidance on possible approaches to their resolution.” Also, statisticians may be active in many other domains than official statistics, *e.g.*, medical research, and many other activities than research, *e.g.*, teaching and consultancy. It goes without saying that professional codes may have to pay attention to these other domains and activities as well. These qualifications are especially relevant as the origins of the theory of informed consent may be traced back to the Nuremberg Trials of the Nazi doctors’ experiments.

### III.1.1 International Statistical Institute

14. The International Statistical Institute Declaration on Professional Ethics was adopted at its Centenary Celebration in 1985 after a seven year preparation by a Committee composed of, amongst others, Roger Jowell (chair), W. Edwards Deming, and Edmund Rapaport. The Declaration contains four sections, devoted to obligations to society, obligations to funders and employers, obligations to colleagues, and obligations to subjects, respectively. Of these, the latter section is almost as extensive as the former three taken together. The section about obligations to subjects is subdivided into six subsections. Of these, two pertain to obtaining informed consent and modifications to informed consent. The combination of these two subsections is longer than any single section of the other three sections. So without knowing any further, one might already conclude that the issue of informed consent is both important and difficult for statisticians. It therefore seems wise to quote these two clauses *in extenso*:

#### 4.2 Obtaining informed consent

Statistical inquiries involving the active participation of human subjects should be based as far as practicable on their freely given informed consent. Even if participation is required by law, it should still be as informed as possible. In voluntary enquiries, subjects should not be under the impression that they are required to participate; they should be aware of their entitlement to refuse at any stage for whatever reason and to withdraw data just supplied. Information that should be likely to affect a subject’s willingness to participate should not be deliberately withheld.

#### 4.3 Modifications to informed consent

On occasions, technical or practical considerations inhibit the achievement of prior informed consent. In these cases, the subjects’ interests should be safeguarded in other ways. For example:

- 1) Respecting rights in observation studies. ...
- 2) Dealing with proxies. ...
- 3) Secondary use of records. ...
- 4) Misleading potential subjects. ...

15. Several arguments deserve attention. In Clause 4.2, the principle of obtaining informed consent is restricted to active participation of human subjects. In addition to Clause 4.2 the ISI Declaration contains a list of twelve “items of information that are likely to be *material* to a subject’s willingness to participate. ... In selecting from this list, the statistician should consider not only those items that he or she regards as material, but those which the potential subject is likely to regard as such.” The ISI Declaration, however, also states that “One way of avoiding inconvenience to potential subjects is to make more use of available data instead of embarking on a new inquiry. For instance, by making

greater use of administrative records, or by linking records, information about society may be produced that would otherwise have to be collected afresh.” The text following Clause 4.3 explains that “A serious problem arises for statisticians when methodological requirements conflict with the requirement of informed consent.”

### **III.1.2 American Statistical Association**

16. The ASA Ethical Guidelines for Statistical Practice, published by the American Statistical Association in 1989, cover less than two pages. They do not mention the concept of informed consent explicitly. Instead, the ASA Guidelines do call upon statisticians to “inform each potential respondent about the general nature and sponsorship of the inquiry and the intended use of the data.”

17. The Guidelines are being revised at present and the draft guidelines are substantially longer. One (out of eight) general topic area(s) is devoted to responsibilities to research subjects, including survey respondents and persons and organisations involved with administrative records, as well as subjects of physically or psychologically invasive research. But the new clauses do not explicitly make mention of informed consent either. Extensive thought has been given to the right choice and phrasing of the new guidelines, as is evident from the various discussions at the Joint Statistical Meetings and in the various ASA publications, including its website. In fact, the notion of informed consent has been deleted from an earlier draft. It was argued that statistical practitioners can not personally accept the responsibility to “obtain” informed consent of human research subjects. The argument was advanced in the context of medical statistics where often statisticians are responsible for only a part of the larger research design. It was accepted, however, for the field of statistics in general.

### **III.1.3 American Sociological Association**

18. The ASA Code of Ethics, approved by the American Sociological Association membership in 1997, is quite elaborate and explicit. One out of twenty items concerns confidentiality, another item is explicitly devoted to informed consent. Both items take several hundreds of words. The two together make up a quarter of the full code. Informed consent is defined as “a basic ethical tenet of scientific research on human populations. Sociologists do not involve a human being as a subject in research without the informed consent of the subject or the subject’s legally authorised representative, except as otherwise specified in this Code.” Sociologists may seek waivers when the risks for research subjects are minimal or the research could not practicably be carried out were informed consent to be required. But such waivers require approval from institutional review boards or, in their absence, from another authoritative body with expertise on the ethics of research. Under such circumstances the confidentiality of any personally identifiable information must be maintained at any rate. In cases of doubt, consultation of institutional review boards is thought to be required, and confidentiality of any personally identifiable information is to be maintained *a fortiori*. In obtaining informed consent sociologists are to use language that is understandable to and respectful of research participants or their legal representatives. Also, they are to keep records regarding said consent.

### **III.2.1 United Nations**

19. In 1992 the United Nations Economic Commission for Europe accepted the Fundamental Principles of Official Statistics. The basis was laid by the Conference of European Statisticians. It attempted codification on behalf of the new democracies and their national statistical institutes in Central and Eastern Europe. Later, the United Nations Statistical Commission endorsed the principles. In its considerations the Resolution holding the Fundamental Principles recalls the ISI Declaration on Professional Ethics. Nevertheless, it does not mention the concept of informed consent explicitly. The two principles that come closest, are the following:

- Data for statistical purposes may be drawn from all types of sources, be they statistical surveys or administrative records. Statistical agencies are to choose the source with regard to quality, timeliness, costs and the burden on respondents.
- Individual data collected by statistical agencies for statistical compilation, whether they refer to natural or legal persons, are to be strictly confidential and used exclusively for statistical purposes.

### III.2.2 Council of Europe

20. The Council of Europe has accepted a Recommendation on the protection of personal data collected and processed for statistical purposes, in 1997. In its considerations the Council shows itself explicitly aware of the necessity to balance the public need to know versus the individual right to privacy, but also of the advances in statistical methodology and information and communication technology over the past decade.

21. The Recommendation itself gives a number of requirements that should be met when collecting personal data for statistical purposes. When using administrative data for statistical purposes, proper notification should be given, unless this would be impossible or require unreasonable effort or provisions have been made in national law. If consent by the data subject is required, it should be free, express, and informed. The fact that the survey serves a legitimate goal may not outweigh the obligation to ask for consent, unless a serious public interest justifies such an exception.

22. Specific provisions have been taken to give legitimacy to proxy interviews and non-response analysis.<sup>7</sup> Personal data may be collected from a member of the household rather than the data subject himself, depending on the nature of the survey. Data on data subjects that are necessary for the preparation and execution of the survey and data about non-response may be used, but only to assure the representative character of the survey.

23. Further provisions relate to the transmission of personal data for statistical purposes, i.e., the release of microdata. A final feature is the call upon professions and institutions that produce statistics to establish and publish a code of ethics. One or more independent authorities should be given the task to ensure national upholding of the principles embodied in the Recommendation of the Council of Europe.

### III.2.3 European Union

24. Three pieces of European Union legislation bear on the production of official statistics, statistical confidentiality, and the protection of personal data. The oldest of these, Regulation 1588/90, made possible the transmission of confidential statistical data from the national statistical institutes to Eurostat, the EU statistical office. Under certain conditions it put aside national statistical confidentiality legislation that used to prohibit the release of national, confidential data to Eurostat. The text of the regulation does not mention data-subjects: Eurostat does not have a direct relation with them. The debate about this regulation has triggered many measures, including funding research into confidentiality methodology and software, however.

25. The EU Data Protection Directive (95/46/EC) went into force in 1995. It attempts to harmonise national legislation within the European Union so as to make it easier to exchange personal data without violating privacy constraints. One feature is the shift of attention from registrations to data.<sup>8</sup> In the Directive personal data are defined as any information pertaining to an identified or

---

<sup>7</sup> These two activities should otherwise be considered to run counter to a strict application of the informed consent principle.

<sup>8</sup> As a consequence the Netherlands' privacy authority, the *Registratiekamer*, will have to change its name.

identifiable natural person. As identifiable is considered a person who can be identified directly or indirectly, specifically on account of an identification number or one or more specific elements characteristic for his or her physical, physiological, psychical, economic, cultural or social identity. The Directive has given the member-states three years time to adapt national legislation. In the Netherlands privacy legislation is relatively recent and does not need to be revised and adapted very extensively. Nevertheless, the deadline has not been met, as seems to be the case throughout the Union.

26. The European Statistical Law, Regulation 322/97/EC, lays down the principles of the European statistical system. In the considerations ample tribute is paid to some of the texts mentioned above. The Statistical Law itself mentions important principal foundations for the European statistical system, such as statistical confidentiality and transparency. The latter principle implies the right of respondents to information about the legal foundation, the objectives for which data are requested, and the protective measures taken. The principle of informed consent is invoked but only when it comes to the use of data collected for any other but the statistical purposes for which they were collected.

### III.3.1 Netherlands statistical legislation

27. The recent (1996) Netherlands' statistical law gives Statistics Netherlands in principle the right of access to administrative data throughout government. In particular it authorises Statistics Netherlands to use the social-fiscal identification number for statistical purposes. Such legal authorisation is required for legitimate use of this personal identification number when processing personal data. Since the number is contained in many relevant administrative files, it greatly enhances the proficiency of matching them.

28. The law also states quite explicitly that all data with Statistics Netherlands are to be kept confidential and shall not be used for any other than statistical purposes. Identifiable individual data may be released only if there is grounded reason to assume that those concerned do not object.<sup>9</sup> The authority to release microdata is given to Statistics Netherlands in explicit deviation from the principle of statistical confidentiality. In the implementation several lines are followed simultaneously to ensure that both the data and the settings are as safe as possible (Kooiman, Nobel and Willenborg, 1999). Apart from the principle of statistical confidentiality and the grounded reason mentioned above, the Netherlands' statistical law does not refer specifically to the relation of Statistics Netherlands with its data subjects.<sup>10</sup>

### III.3.2 Netherlands privacy legislation

29. The Netherlands' Privacy Law, WPR (1988), belongs to the second generation of European privacy legislation. It is largely procedural. Its basic principles state that for processing personal data a legitimate goal is required; that the actual use of the data has to be compatible with the purposes for which they had been collected; and that the processing itself has to be transparent for those concerned. The law makes several exemptions for processing on behalf of statistical and research purposes. In fact, the legislator has deemed the social utility of statistics and research as outweighing individual interests. Data subjects have no right to notification, access, and correction of data about them that are in use for statistical and research purposes only. Holders of administrative files are entitled (but not compelled) to transmit these files to Statistics Netherlands for statistical purposes. In order to implement this entitlement, the privacy statute governing the administrative data has to make mention of such transmission, however. If Statistics Netherlands succeeds in returning added value (*i.e.*, aggregate information) to the original holders, permission to use administrative data is often easier to arrange.

<sup>9</sup> This is not necessarily the same as the principle of informed consent.

<sup>10</sup> For social surveys, that is. For business surveys, the 1996 law has several statements on response burden. Also, there is a specific 1936 Law on Economic Statistics with more specific statistical confidentiality statements.

30. The position of the Netherlands *Registratiekamer*, the national privacy authority, towards Statistics Netherlands is formal. The law entitles the statisticians to use administrative data including the social fiscal identification number and to release microdata. Statistics Netherlands is considered a model in terms of sealing and securing its data processing. Beyond that, it seems that the *Registratiekamer* has no particular or strong feelings about (our feelings about) our relations with respondents.

31. So far fiscal data have not been matched with survey data because Statistics Netherlands deems it necessary to inform its respondents about such matching. The fiscal authorities do not oppose matching itself but do object to informing respondents. They fear that as a consequence of such information the public might distrust fiscal confidentiality or even might change its fiscal reporting behaviour. The royal way is to provide a stronger political backing for the proposed matching.

#### IV. CASE STUDIES AND PILOT SURVEYS

32. A few case studies have been done and are in progress at Statistics Netherlands to explore the implications of informed consent to match registration-based data with survey responses. Some results will be reported in this section.

33. In the first stage of research, in 1995, a number of focus group sessions have been held. One important conclusion was that respondents had objections against matching because they had the feeling of losing control of their data. They preferred that permission was asked in ways as concrete as possible. The second conclusion was that asking for consent for matching in combination with the usual pledge of anonymity led to confusion amongst the respondents. The third conclusion was that resistance amongst respondents might turn out to be subject matter related. On the basis of these conclusions it was decided that in the second stage concrete register data (fiscal income data, municipal housing data) were to be mentioned. Different data collection procedures (in depth interviews, CAPI, CATI) were to be used. The framework of the general life situation survey (POLS) was used as the carrier.

34. Two experienced interviewers asked fourteen test respondents for their permission to retrieve income data from the fiscal authorities, in a casual fashion at the end of the interview. Five respondents gave permission to use their fiscal income data and did not care to respond to the income questions. Six respondents had some objections, however.<sup>11</sup> One respondent needed his parents' permission. And two had no objection to matching records but preferred after all to answer the questionnaire for the income variables. In the end the income questionnaire was used instead of fiscal data for five respondents. All in all, a majority of respondents were willing to contribute income data. But it took quite some effort and they wanted themselves to determine the exact source of such data. Less than half consented to giving access to fiscal records.

35. In a 1996 fieldwork experiment among 400 respondents the difference between CAPI and CATI did hardly turn out to be significant. Two out of three respondents granted immediate access to their housing data resorting with the municipal government. Another twenty percent consented after some hesitation. Only ten percent refused. Several of them volunteered very clear statements about their motives, usually in the form of objections against the municipal government and administration. One significant observation came from respondents in rental homes who thought their landlord's permission more appropriate. The main finding suggests that the permission to retrieve and match administrative

---

<sup>11</sup> One respondent was taken by surprise, one fell silent, a third became red arguing that she could choose not to talk about these data that no one had anything to do with; the fourth wondered if this were allowed and preferred to answer the questions. The fifth preferred to answer himself. The sixth refused without further comment.



data may depend on the sensitivity of the issue, where income is considered a much more private and delicate issue than housing.

36. In the CATI part respondents were also asked if they would be willing to participate in subsequent surveys. People who did not give consent to retrieve and use their administrative housing data, were also less willing to participate in subsequent surveys. Furthermore, those who refused to respond to the income question were also inclined to deny access to the administrative data.

## V. CONCLUSIONS

37. The approach of data subjects by statistical offices has become more complicated at the same pace as the statistical production process itself. Legislation and ethics do not always provide guidelines for new dilemmas. What then are my conclusions from the above? In my view the notion of “informed consent” is neither a buzzword nor a panacea. It is an appeal to take serious the relation with our data subjects while keeping in mind our objective of providing useful and valid statistical information. In order to reach this objective our data are our most precious asset. If only for this reason - we will have to produce statistics in the future - we must take our data subjects serious.

But in official statistics informed consent is not a moral imperative that is to be followed under all circumstances and irrespective of consequences. The availability of official statistics for society is a public good without which government and the market could not function properly. In the past and in other countries than the Netherlands a response obligation helped provide this information. In the Netherlands such an obligation is in existence since 1917 for economic statistics. Administrative data by themselves are used for statistical purposes without data subjects being aware of such operations.

38. Statisticians must do what they promise to do. They must not promise more than they can do. And they must perceive their promises through the eyes of their data subjects. If they promise to guarantee confidentiality, they must be sure that they are in a position to do so. Both legal and methodological reasons may be reason for a more cautious approach. They must be sure that administrative authorities have no title at all to claim their data. And they must be sure, that data are absolutely non-identifiable, which is a very high, if not impossible, standard. And if they promise that the data will be used for statistical purposes only, they must be sure that their data subjects have the same understanding of this formula as they do. If they intend to cover the release of microdata to research institutes under this formula, this intention must be clear to respondents. Legalistic defences like “but we never promised that the microdata would not leave our office” appear to be objectionable from an ethical point of view.

39. To proceed legitimately, statisticians must comply with all legal provisions about data confidentiality and privacy. They should receive data only from those who are authorised to respond or transmit data. In particular, they have to comply with the conditions posed by those in charge of the administrative files. If these do not want to have respondents informed about matches with their data without political backing, they shall have to provide such political backing. Statisticians should not ask for more data than is warranted for statistical purposes. Identifiers should be separated as soon as possible and stored, if necessary at all, during as short a period as possible, outside the direct authority of those working with the substantive microdata. They should also take the technical measures that are necessary to keep their data confidential, such as working on dedicated and isolated networks.

40. If statisticians seek informed consent from respondents, they must do so seriously. Information must be fair and full.<sup>12</sup> If they inform respondents about their intention to match survey records with administrative data, they must name these administrations. They must include at least those administrations that are of greatest interest to them. Specifically, fiscal files may not remain

---

<sup>12</sup> This is not the same as naive. A two-tier approach with details available in a brochure following an attractive advance letter of invitation may help prevent abhorring potential respondents.

unmentioned (covered by “et cetera”), if such matching is intended. Consent must be asked and given explicitly. If they request consent from respondents for such matching, they must take their answers seriously. If respondents prefer answering standard survey questions, they should be given the opportunity. And if they deny access to administrative records, matching should not be performed behind their back after they have been asked for permission.

41. But we should not necessarily stop at the position of informed consent if this would lead to qualitative defects in statistical results. Statistical inference could be seriously hampered if, *e.g.*, for one part of the respondents fiscal income data would be available, for another part survey income data, for yet a third part both kinds, and for the last part no income data at all. One solution would be to base the statistical system as a whole on administrative data only and not to burden respondents with a request for informed consent. Matching with survey data could occasionally be done in addition to this system.

42. A more radical solution might be to seek for a formal and explicit authorisation to match administrative and survey data. Parliament or the national privacy authority might conclude that the threats to confidentiality and privacy of such matches are minimal and the advantages of the resulting information outweigh them by far. Furthermore such political backing may be constrained by substantive and/or procedural conditions. For example, one such condition might make the intended matching dependent on explicit authorisation by the Central Commission for Statistics, the *Registratiekamer*, or some panel of civilian representatives. At any rate, if such political support can be brought about, notification rather than informed consent is the way to approach the data-subjects.

## Bibliography

American Sociological Association 1998. *Code of ethics*. <http://www.asanet.org/ecoderev.htm>

American Statistical Association 1989. Ethical Guidelines for Statistical Practice. *AMSTAT News* 1989 (154):24-25. <http://www.tcnj.edu/~asaethic/asagui.html>

Beauchamp, T.L., R.R. Faden, R.J. Wallace, and L. Walters, eds. 1982. *Ethical Issues in Social Science Research*. Baltimore: Johns Hopkins University Press

Bochove, C.A. van, 1996. From assembly line to electronic highway junction: a twin-track transformation of the statistical process. *Netherlands Official Statistics* 11(2): 5-36.

Council of Europe 1997. *Recommendation No. R(97)18 and explanatory memorandum of the Committee of Ministers to Member states concerning the protection of personal data collected and processed for statistical purposes*. [http://www.coe.fr/DataProtection/rec/r\(97\)18e.htm](http://www.coe.fr/DataProtection/rec/r(97)18e.htm)

Duncan, G.T., T.B. Jabine, and V.A. de Wolf, eds. Committee on National Statistics, Commission on Behavioural Sciences and Education, National Research Council, and the Social Science Research Council. 1993. *Private lives and public policies – Confidentiality and accessibility of government statistics*. Washington, D.C.: National Academy Press.

Economic Commission for Europe 1992. *The fundamental principles of official statistics in the region of the Economic Commission for Europe*. <http://www.unece.org/stats/archive/docs.fp.e.htm>

Ellenberg, J.H. 1983. Ethical guidelines for statistical practice: a historical perspective. *The American Statistician*. 37(1):1-5.

European Union 1990. *Council Regulation (Euratom, EEC) No 1588/90 of 11 June 1990 on the transmission of data subject to the Statistical Office of the European Communities*. <http://europa.eu.int/eur-lex/en/390R1588.html>

European Union 1995. *Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data.* [http://europa.eu.int/eur-lex/en/lif/dat/en\\_395L0046.html](http://europa.eu.int/eur-lex/en/lif/dat/en_395L0046.html)

European Union 1997. *Council Regulation (EC) No 322/97 of 17 February 1997 on Community Statistics.* [http://europa.eu.int/eur-lex/en/lif/dat/en\\_397R0322.html](http://europa.eu.int/eur-lex/en/lif/dat/en_397R0322.html)

Fienberg, S.E., M.E. Martin, and M.L. Straf, eds. Committee on National Statistics, Commission on Behavioural and Social Sciences and Education. 1985. *Sharing Research Data*. Washington, D.C.: National Academic Press

International Statistical Institute 1985. *ISI Declaration on professional ethics.*

<http://www.cbs.nl/isi/ethics.htm>

Jabine, T.B. 1986. Selected guidelines for notification to survey participants. *American Statistical Association 1986 Proceedings of the Section on Survey Research Methods*. Washington, D.C.: American Statistical Association: 1-5.

Jowell, R. 1981. A professional code for statisticians? Some ethical and technical conflicts. *Bulletin of the International Statistical Institute* (Proceedings of the 43rd session) 66:273-276.

Knott, J. Administrative records use in the 1995 Census test. B. Jamerson and W. Alvey, eds. *Turning Administrative Systems into Information Systems: 1994*. Washington, D.C.: Statistics of Income, Internal Revenue Service: 57-68.

Kooiman, P., and J.R. Nobel. 1997. Vraag nooit naar de bekende weg: statistisch beleidsonderzoek met registraties. (Statistical policy research using administrative data.) In: P. Ester, J.L.A. Geurts, and M.J.M. Vermeulen, eds. *De makers van de toekomst*. (The makers of the future.) Tilburg: Tilburg University Press.

Kooiman, P., J.R. Nobel, and L. Willenborg 1999. Statistical data protection at Statistics Netherlands. *Netherlands Official Statistics* 14: to appear.

Kordes, F.G. 1997. *Privacy-wetgeving en het gebruik van persoonsgegevens voor wetenschappelijke en statistische doeleinden*. (Privacy legislation and the use of personal data for research and statistical purposes.) The Hague: Ministry of Education, Culture and Sciences.

Mugge, R.H. 1993. Informed consent in U.S. government surveys. *Journal of Official Statistics*. 9(2): 345-360.

Presser, S. 1994. Presidential address – informed consent and confidentiality in survey research. *Public Opinion Quarterly* 58(4):446-459.

Singer, E. 1993. Informed consent and survey response: a summary of the empirical literature. *Journal of Official Statistics*. 9(2):361- 375.

Wet, 1996. *Wet van 18 april 1996, houdende een instellingsregeling voor het Centraal bureau en de Centrale commissie voor de statistiek*. (Netherlands' Statistics Law). Staatsblad 1996: 258.

WPR, 1988. *Wet van 28 december 1988, houdende regels ter bescherming van de persoonlijke levenssfeer in verband met persoonsregistraties*. (Netherlands' Privacy Law.) Staatsblad 1988: 665.

<http://www.unimaas.nl/~privacy/wpr.htm>